

E-CONOM

Online tudományos folyóirat
Online Scientific Journal

Tanulmányok a gazdaság- és társadalomtudományok területéről
Studies on the Economic and Social Sciences



E-CONOM

Online tudományos folyóirat | Online Scientific Journal

Főszerkesztő | Editor-in-Chief
KOLOSZÁR László

Kiadja | Publisher
Soproni Egyetem Kiadó |
University of Sopron Press

A szerkesztőség címe | Address
9400 Sopron, Erzsébet u. 9., Hungary
e-conom@uni-sopron.hu

A kiadó címe | Publisher's Address
9400 Sopron, Bajcsy-Zs. u. 4., Hungary

Szerkesztőbizottság | Editorial Board
CZEGLÉDY Tamás
HOSCHEK Mónika
JANKÓ Ferenc
SZÓKA Károly

Tanácsadó Testület | Advisory Board
BÁGER Gusztáv
BLAHÓ András
FÁBIÁN Attila
FARKAS Péter
GILÁNYI Zsolt
KOVÁCS Árpád
LIGETI Zsombor
POGÁTSA Zoltán
SZÉKELY Csaba

Technikai szerkesztő | Technical Editor
TAKÁCS Eszter

A szerkesztőség munkatársa | Editorial Assistant
DURGULA Judit

ISSN 2063-644X



Tartalomjegyzék | Table of Contents

ERB Franciska – GÖZ András

Feleletválasztós tesztek kiértékelésének módszerei és a kiértékelést támogató alkalmazás fejlesztése

Methods for analyzing simple choice tests and developing an application which supports the evaluation3

UZONYI Antal

Hajdúböszörmény helyzetének elemzése, valamint az agrár- és vidékfejlesztési támogatások értékelése

Analysis of the situation of Hajdúböszörmény, and evaluation of agricultural and rural development subsidies15

BURKA Nikoletta Zsuzsanna

A hidegen sajtolt tökmagolaj mint helyi termék szerepe a térségi fejlődési folyamatokban

Role of the local production of the cold-pressed pumpkin seed oil in the local development processes.....32

LIPTÁK Lilla

A közösségi tagság hatása a vegán fogyasztók magatartására

The effect of community membership on vegan consumers' behavior46

ILYÉS Noémi

Láthatatlan turisták Budapest árnyékában

Invisible tourists under the shadow of Budapest56

MOLICH Kitti – STEFÁN Ádám

Check-in a jövőbe – robotizálás a szállodai recepción

Check-In To The Future – Robotisation At Hotel Reception66

ERB FRANCISKA
GŐZ ANDRÁS¹

Feleletválasztós tesztek kiértékelésének módszerei és a kiértékelést támogató alkalmazás fejlesztése^{2 3}

A számonkérés egyik alapvető, gyakran használt formája a teszttíratás. A megfelelően kialakított tesztekkel könnyebb, gyorsabb a hallgatók tudását felmérni, ugyanis minimális javítási időt igényelnek, főleg, ha nem papíron, hanem programban meghatározott feltételek mellett adnak számot a diákok az elsajátított tudásról. Kutatásunk központjában a Budapesti Corvinus Egyetem Információrendszerek Tanszékének egyik meghatározó számonkérési tesztrendszer és tesztkörnyezete állt. Az oktatóknak eddig nem volt lehetőségük arra, hogy visszajelzést kapjanak a tesztors hatékonyságáról, és hogy felismerjék a mára elavult tesztkérdéseket. Célunk mindazon kérdések azonosítása volt, melyek nem megfelelően mérték fel vagy differenciálták a hallgatók tudását, majd a kapott eredménnyel iránymutatást kívántunk adni mind a kérdések javításához, mind a kapcsolódó tananyagok fejlesztéséhez. További célunk egy olyan alkalmazás fejlesztése volt az oktatók számára, amely elemzi a vizsgasor kérdéseit, elvégzi a szükséges számításokat és azonnali visszajelzést ad az egyes vizsgák eredményeiről.

Kulcsszavak: véletlen együttható, regresszió, korreláció, statisztikai módszerek
JEL-kódok: C250

Methods for analyzing simple choice tests and developing an application which supports the evaluation

Tests are frequently applied for measuring students' knowledge. With adequately prepared tests measuring knowledge acquired is much easier. After collecting a certain amount of test results, it's recommended to analyse test questions in the light of the results, to point out mistakes and imperfections of the test repository. In our work we analysed test questions and results of the Information Management course of the Business Information Systems training program at Corvinus University of Budapest. A crucial element of course development is to provide feedback to the teacher concerning the efficiency of the applied tests and enable the recognition of wrong questions. Our goals were to identify questions that can not measure the knowledge of students perfectly and offer guidelines based on the results for improving the quality of tests and the related knowledge base. The other objective of our research was to develop an application for teachers, which helps to analyse test results and the quality of test questions and provides immediately feedback this way.

Keywords: Random Coefficient, Regression, Correlation, Statistical Methods
JEL Codes: C250

¹ A szerzők a Budapesti Corvinus Egyetem Gazdálkodástudományi Karának Gazdaságinformatika szakon végzett alapszakos hallgatói (erbfrenciska@gmail.com, gozandris@gmail.com).

² A tanulmány a XXXIII. Országos Tudományos Diákköri Konferencia Gazdálkodástudományi Kar Információmenedzsment szekciójának Információ-gazdaságtan, Gazdasági informatika és e-business II. Tagozatában első helyezést elért dolgozat alapján készült. Az OTDK-pályamunka konzulense dr. Vas Réka Franciska egyetemi docens.

³ A tanulmány az OTDT Közgazdaságtudományi Szakmai Bizottsága, valamint a XXXIII. OTDK Közgazdaságtudományi Szekciójának szervezője, a Széchenyi István Egyetem gondozásában kiadott "Litera Oeconomiae - Válogatás a XXXIII. OTDK Közgazdaságtudományi Szekció helyezést elért pályamunkáiból" című tanulmánykötetben (2017) is megjelent.

Bevezetés, célok

A Budapesti Corvinus Egyetem Információrendszerek Tanszékén számos olyan (főként gyakorlati jeggyel záruló) tantárgy található, amelynél az oktatók tesztek segítségével mérik és értékelik a félév során (esetenként több alkalommal) a hallgatók megszerzett tudását. Elemzéseink során e tárgyak számonkérési rendszereit vizsgáltuk meg alaposabban, mivel a tesztsorozatok számos frissítést, változtatást igényeltek már. Az oktatóknak nem volt lehetőségük arra, hogy a teszt megíratása után rövid összefoglalást, visszajelzést kapjanak a teszt sor hatékonyságáról. A kérdéskatalógusokban ráadásul akadtak olyan kérdések is, amelyek mára elavulttá váltak.

Elsődleges célunknak tekintettük mindazon kérdések azonosítását tantárgyanként, amelyek már nem megfelelően mérik fel a hallgatók tudását, nem képesek megfelelően differenciálni a hallgatókat tudásuk alapján, majd a kapott eredménnyel iránymutatást kívántunk adni mind a kérdések javításához, mind a kapcsolódó tananyagok fejlesztéséhez, rámutatva azokra a témákra, amelyekre több időt érdemes szánni, hogy jobban megérthessék azokat a hallgatók.

Továbbá kutatómunkánk célja volt az azonos tesztkörnyezetben futó tesztek és feladatsorok részletes elemzéséhez egy olyan keretrendszert biztosítani, amely rámutat a feladatsorok esetleges hibáira, hiányosságaira. Ezzel növelhető annak a hatékonysága és eredményessége, hogy az egyetemi oktatók a rendelkezésükre álló korlátozott időben adott időszak után elemezni tudják számonkérési rendszerüket, megvizsgálják annak hatékonyságát és eredményességét, ami elengedhetetlen a sikeres hallgatói és oktatói munka és teljesítmény értékeléséhez.

Számos módszer és megoldás áll rendelkezésre a tesztsorozatok, azon belül a kérdések és válaszok elemzésére, illetve a hallgatói teljesítmény értékelésére a szakirodalomban. A rendelkezésre álló megoldások kiválasztása és meghatározása után egy adott környezetben gyors elemzési lehetőséget lehet biztosítani az oktatók számára, így a kapott eredmények értelmezése után a tesztsorozat fejlesztésére és javítására is marad érdemi idejük és lehetőségük.

Felállított keretrendszerünk segítségével megszerezni tudtuk és áttekinthetővé tudtuk tenni a számításainkat, amely így a jövőben hozzájárulhat az egyes kurzusok során zajló színvonalas oktatás fenntartásához és elősegítheti nem csak a teszt sorok, hanem a tananyagok folytonos ellenőrzését megújulását is.

Az elemzés során segítségünkre volt egy olyan általunk fejlesztett alkalmazás, amely képes a kiindulási adatokat kezelhető formába rendezni amellyel, hogy az oktatók számára azonnal használható visszajelzést is ad az egyes tesztek, teszt sorok hatékonyságáról, eredményességéről.

Témaválasztásunk oka az értékteremtés, hiszen a megfelelő iránymutatással hozzájárulunk mind az oktatók, mind a hallgatók sikerességéhez. A gazdaságinformatikus képzésünkben szereplő Informatikai rendszerek fejlesztése című tárgy keretein belül hasonlóan nagy adatbázisokkal foglalkoztunk, így a dolgozat egy részét, a programalkotást a tárgy keretein belül megszerzett tudásra építettük, kipróbálva magunkat új problémák megoldásában, továbbá érdekesnek tartottuk, hogy többek között a saját szakunkon tanult tárgyak eredményeivel is dolgozhatunk, és hozzájárulhatunk a következő évfolyamok tanításának fejlesztéséhez.

Összegezve azt mondhatjuk, hogy a rendelkezésre álló különböző tantárgyi adatok részletes elemzése után a dolgozatban választ adtunk arra, hogy az egyes teszt sorok kérdései megfelelően mérik-e a hallgatói teljesítményt, melyek a legnehezebb és legkönnyebb vagy a legrosszabb és legjobb kérdések a különböző mérések alapján a feladatsorokban, mely kérdések váltak elavulttá a vizsgált időszakok alatt, mit mutatnak az éves tendenciák, van-e összefüggés az éves átlageredmények között és hogyan tudjuk a kapott eredményeket megfelelően felhasználni úgy, hogy a tesztsorozatok megújítása, frissítése valóban rendszeresen megvalósuljon. Amellett, hogy a kiválasztott tantárgyakat egyesével a saját környezetükben elemeztük, kíváncsiak voltunk arra is, hogy milyen hasonlóságok vagy éppen különbségek fi-

gyelhetők meg, ha egymás mellé állítjuk a fontosabb mutatók eredményeit és a tantárgyakat egymáshoz viszonyítva is megvizsgáljuk.

A kutatás eredményeként egy olyan megoldás vált elérhetővé az oktatók számára, amely gyors és részletes elemzést nyújt adott kiinduló adatok beolvasását követően a számonkérési tesztsorok kérdéseinek, válaszainak és a hallgatók eredményeinek megvizsgálására, adott feltételek mellett meghatározott releváns módszerek és megoldások gyűjteményeként.

Az elemzési szempontok egységes környezetben történő elhelyezése pedig lehetőséget biztosít az oktatóknak a kapott eredmények analizálására és javítására, egyszerűbbé téve az újítási szándékok megfogalmazását és bevezetését. Mindez hozzájárulhat a hallgatói teljesítmény pontos méréséhez és az adott tárgy színvonalas számonkérési rendszerének fenntartásához, melyek az oktatói és hallgatói sikerességet és kiválóságot erősítenék az egyetemen.

Módszertan

A feleletválasztós tesztek kiértékelését támogató keretrendszerünk kialakításánál statisztikai, matematikai módszertanok, módszerek és az item analízis eszköztárát használtuk fel.

A tesztek értékelését és fejlesztését az teszi lehetővé, hogy kiforrott matematikai, statisztikai elméletekkel rendelkezünk, melyekből olyan mutatókat vezethetünk le, amelyek alkalmasak a teszt hatékonyságának, jóságának mérésére (Nagy 1975; Horváth 1991, 1993; Murphy és Davidshofer 1994). Ugyanazokat a teszteket igen különböző és nagyszámú tanuló-csoport is megoldhatja, ezért fontos, hogy a megfelelő feladatokat egy adott idő után statisztikai elemzések alá vessük, felmérve azok tulajdonságait.

A statisztika alapkérdései egyrészt az általános tendenciák kimutatására, másrészt a középértékektől való eltérések feltárására, harmadrészt az értékek közötti összefüggések vizsgálatára irányulnak. A leíró statisztika segítségével az adatok könnyen átlátható, kezelhető alakra hozhatók, módszerei alkalmazásával pedig leírhatjuk az egyes változókat, illetve a változók közötti kapcsolatokat. Mérőszámok a gyakoriság, középértékek, szóródás. E mutatók kiszámítása után kapott eredmények mindig csak a vizsgált mintára érvényesek. A pedagógiai kutatásoknál azonban az a cél, hogy egy mintán nyert adatokból minél nagyobb populációra tudjunk következtetéseket levonni. E következtetések megadásához használható fel a statisztikai hipotézisvizsgálat különböző próbák alkalmazásával, illetve a variancia-, korreláció- és regresszióanalízis (Kontra 2011).

Az úgynevezett item mutatók, a tesztet felépítő legkisebb, még önállóan megítélhető egységekkel, az itemekkel (kérdésekkel) foglalkoznak, és fontos szerepet töltenek be a tesztek értékelésénél, fejlesztésénél. Az item nehézsége úgy határozható meg, hogy a helyes, jó megoldások számát elosztjuk az összes megoldás számával. Minél közelebb van az item nehézsége az 1-hez, annál többen válaszolták meg helyesen az adott kérdést és minél közelebb van ez az érték a 0-hoz annál kevesebben adtak jó választ. E mutató másik gyakran használt neve a P-mutató. Általában a 0,9-es érték feletti itemeket a roppant könnyű, míg a 0,2-es érték alattiakat a borzasztóan nehéz kategóriába sorolják. Az itemek szóródása igen fontos információ az oktatás hatékonysága szempontjából és összefügg az item-nehézséggel. A szórás az átlagtól való átlagos eltérést számszerűsíti és összefügg a nehézséggel. Jelentős mutató, mert sok matematikai-statisztikai módszer épül a szórásra, illetve a varianciára (szórásnégyzetre) (Csapó, 2002).

Az elkülönítés-mutató az item azon tulajdonságát adja meg, hogy mennyire képes differenciálni a hallgatókat az alapján, hogy mennyire jól tudják a vizsgálandó tananyagot. Korábban különböző manuális, kézi, hagyományos számolásokat alkalmaztak az értékek meghatározására, mint például a diákok alacsonyabb és magasabb kategóriákra bontása tudásuk alapján, hogy így hasonlítsák össze az itemekre érkezett válaszokat az elért összpontszámokkal. A számítógépes elemzéssel azonban pontosabban meg tudjuk határozni az egyes itemek elkülö-

nító erejét, mert a kalkuláció figyelembe veszi minden diák minden válaszát a kirívóan alacsony és magas pontszámokkal rendelkező hallgatók eredményeihez képest (Csapó, 2002).

A determinációs mutató ($d[i]$) az adott item teszten belüli befolyásoló képességét határozza meg. A mutatószámot az item találati arány és összpontszám korrelációs együttható segítségével határozzuk meg. A legmagasabb determinációs értékkel rendelkező itemnek van a legnagyobb hatása az összpontszám alakulására, azaz a legnagyobb differenciáló ereje, determinációs hatása.

Az item értékeinek lineáris transzformációi segítségével az adott item átlagának és szórásnak kívánt értékekre beállítását végezzük. Ezt más néven standardizálásnak nevezzük. Ahhoz, hogy pontos és valós eredményeket kapjunk, bizonyos számítások esetén érdemes megadni becslési intervallumot, konfidencia-intervallumot, hogy kezelni tudjuk az ismeretlen paraméterek értékeit egy előre megadott valószínűség mellett. Gyakran alkalmazott konfidenciaszint-értékek a 90%, 95% vagy a 99% (SAKAPKMM, 2010).

A cronbach-alfa mutató segítségével megtudható, hogy a teszt/mérőeszköz mennyire mér megbízhatóan, azaz mennyire reliabilis. A kutatások alapján a 0,7 feletti értékek elfogadhatók, 0,9 fölötti érték esetén pedig a teszt megbízhatóan mér.

A kiszámított item mutatók értelmezésénél fontos az óvatosságra törekedni. Az item adatok kísérleti adatok. Az adatok egy részét befolyásolhatja a kitöltők száma és tulajdonságaik, az oktató által alkalmazott eljárás, és a véletlen hibák. Ha az itemek ismétlődése lehetséges, akkor statisztikailag rögzítve kell lennie minden adminisztrációs jellemzőnek, mely az adott item megjelenéséhez, előfordulásához kapcsolható (Mehrens és Lehmann, 1973).

Elmélet

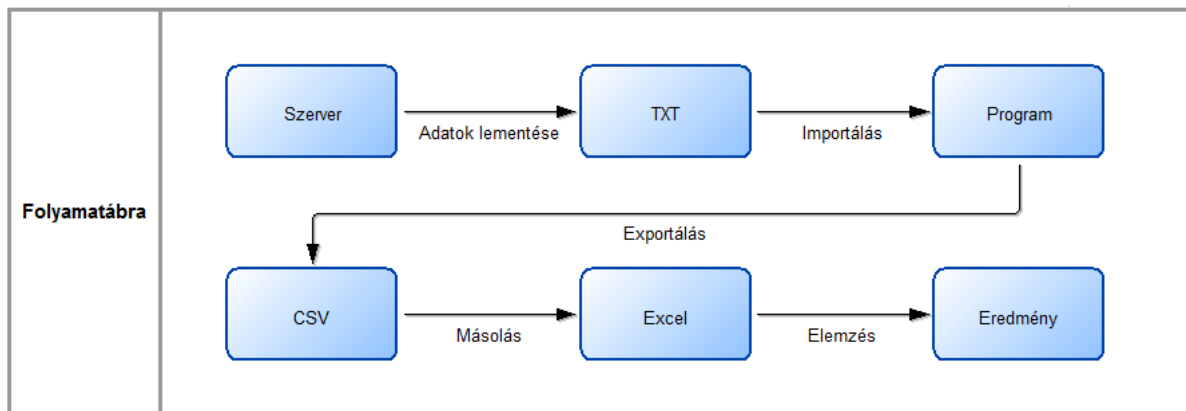
Az elméleti háttér pontos feltárását pedagógiai tanulmányok és módszertanok, különböző egyetemek pedagógus képzéseinek tananyagai és tesztelemzéssel, értékeléssel foglalkozó nemzetközi tudományos cikkek feldolgozása segítette. A pedagógiai mérés egyike a pedagógia legegzezaktabb területeinek. Számos protokollal, szakmai szabállyal és önálló, valódi kismunkált fogalomrendszerrel bír. Képzett szakemberek művelik, akik munkájuk során szükségképpen mérnek, értékelnek és ellenőriznek. Ahhoz, hogy szakszerűen végezzék el tevékenységeiket, szükséges legalább a minimális mérésmetodikai ismeretekkel rendelkezniük. A mérés számos erőforrást igénylő pedagógiai tevékenység. Ahhoz, hogy művelése szakszerű és hatékony legyen, nélkülözhetetlen a megfelelő ismeret, képesség, motiváltság, idő, eszköz, pénz és emberi erőforrás. A mérési rendszereket csak együttműködéssel lehet kiépíteni és működtetni, így az oktatóknak érdemes a csoportos szellemi alkotó technikák alkalmazására törekedni egy új megoldás bevezetése és üzemeltetése során (SAKAPKMM 2010).

A mérés során meghatározott szabály vagy szabályok mentén emberekhez, jelenségekhez, tárgyakhoz, dolgokhoz egy konkrét számot, értéket rendelünk. Az egyik leggyakrabban és legszélesebb körben elterjedt és alkalmazott pedagógiai mérési módszer a tesztelés, tesztíratás. Számos tudományos vizsgálat bizonyította már, hogy hitelesebb eredményt kapunk akkor, ha a tudásszintmérés technikái közül választva a tanulmányi teljesítményeket tesztekkel mérjük, szemben a tanárok értékítéletén alapuló olykor kevésbé objektív osztályzással. A tesztek a pszichológiában és a pedagógiában is fontos szerepet töltenek be, olyan mérőeszközök, melyek elemzése és vizsgálata matematikai és statisztikai modellekre épül. A tesztekkel mérhető tudáson azt a teljesítményt értjük, melyet a tanulók a tananyagra épülő, de közvetlen szubjektív tanári értékelést nem igénylő feladatoknál nyújtanak.

Eredmények

Elemzésünk megkezdéséhez szükségünk volt a vizsgált tantárgyak, kurzusok részletes teszt-eredményeire. Az exportált nyers adatok azonban nem álltak rendelkezésre elemzésre alkal-

mas formátumban. Számos adattisztítási és adatformázási feladatot kellett elvégezni a nyers adatokon a legegyszerűbb tesztelemzések elvégzéséhez is. Ebből következik az, hogy egy feldolgozó program megalkotása vált szükségessé, mely képes az adatok gyors és hatékony átalakítására. A programot a Visual Studio 2013-as kiadású alkalmazásában írtuk, C# programozási nyelven. Az adatok feldolgozásának folyamatát az alábbi ábra mutatja be:



1. ábra: Folyamatábra

Forrás: Saját szerkesztés Adonis programmal

Kiinduló adatainkat a kiválasztott tárgyak esetén a hallgatók által megírt tesztek eredményei jelentették. Az elemzések során a 2009 és 2016 közötti Információmenedzsment, a 2014 és 2016 közötti Üzleti informatika és Üzleti intelligencia és a 2015 és 2016 közötti Integrált rendszerek tárgyakból megírt elméleti tesztek eredményeit vizsgáltuk meg részletesen. A nyers adatok negyedéves, illetve féléves bontásban álltak rendelkezésünkre.

Mennyiségileg tárgyaként különböző a tesztkitöltések száma. Információmenedzsment esetén megközelítőleg 111.500, Integrált rendszerek esetén 17.500, Üzleti informatika esetén 47.000, Üzleti intelligencia esetén 22.000, együttesen majdnem 200.000 megválaszolt teszt-kérdés adatával számoltunk. A hallgatók egyes teszteredményeit egy sorban, „|” jellel elválasztva kaptuk meg. A megalkotott programunk első feladata ezen eredmények szétválasztása és tömbbe rendezése volt, hogy további értékelésre, elemzésre alkalmas adatok, adatsorok álljanak a rendelkezésünkre.

Jelenleg az alábbi funkciók érhetők el a programon belül, melyek az adatok feldolgozását, elemzését segítik:

- Teszteredmények beolvasása.
- Megjelenítési és exportálási lehetőség kérdéskódokként a felmerülések számával, találatok számával, és találati aránnyal.
- A kritikusnak vélt (nehezebb és könnyebb) kérdések megjelenítése és exportálása kérdéskódokként a felmerülések számával, találatok számával, találati aránnyal.
- Megjelenítés és listázás kérdéskódokként a helyes válaszok számával, a kérdés válaszaira érkezett jelölések számával.

Az előállított adatsorokat, listákat a program .CSV formátumban exportálta. Ezt a találati arányok alapján történő számításoknál a „Kérdéskód; Eltalálta; Felmerült; Találati arány” formában, az adott kérdés lehetséges válaszaira érkezett jelölésekkel kapcsolatos kalkulációknál pedig a „Kérdéskód; Helyes válasz; 1. Válasz; 2. Válasz; 3. Válasz; 4. Válasz” formában illeszti be a program az új adattáblába kérdéskódonként csoportosítva soronként 1-1 cellába rendezve az értékeket. Az adatokat beolvasás után még szükséges szétbontani a pontosvesszők mentén, hogy megkapjuk a végleges, elemezhető adattáblákat, majd a kapott adatokat .XLSX formátumban elmentve elvégezhetjük a további statisztikai és matematikai elemzéseket.

Egy tárgy adott időszakra vonatkozó teszteredményeinek, átlagainak összehasonlításával felmérhetjük, hogy összességében a tesztet kitöltő hallgatók milyen eredményt nyújtottak. Ezzel képet kaphatunk a hallgatók, tesztek eredményességéről, és megadhatunk egyben egy indikátort is a teljesítmény változására.

Az elemzések során alkalmazott módszerek és eljárások által kapott eredmények tájékoztató jellegűek az adott vizsgasorral, tesztrendszerrel kapcsolatban. Egymás mellé állíthatók a vizsgált tárgyak eredményei bizonyos mutatók alapján, azonban nem szabad megfeledkeznünk arról, hogy a kiválasztott tantárgyak más-más jellemzőkkel bírnak, különböző a tantárgyak tematikája, felépítése és a hallgatók végső érdemjegyét az elméleti zárthelyi dolgozaton kívül a gyakorlati és plusz feladatok során nyújtott teljesítmény is befolyásolja, melyek mérésére jelen dolgozat keretein belül nem volt lehetőségünk.

A felállított keretrendszer számításai, elemzései:

- Negyedéves átlagok;
- Eredmények átlaga évente tárgyanként;
- Eredmények átlaga összesítve tárgyanként;
- Anova tábla – Várható értékek (átlagok);
- Módusz, medián;
- Teszteredmények terjedelme és szélső értékei;
- Eredmények összehasonlítása tárgyanként;
- Kérdések elemzése:
 - Kérdések minősítése;
 - Elkülönítésmutató;
 - Adott válaszokra érkezett jelölések száma;
 - Összefüggés a helyes válasz és a leghosszabb válasz között, illetve a jelölt válasz és a helyes válasz között;
- Tesztkérdések kategorizálása:
 - Fejezetek, témakörök nehézségeinek vizsgálata;
 - Egyenletesség vizsgálata.

Az általunk megalkotott alkalmazás által előállított adatsorok alapján első lépésként a leíró statisztika körébe tartozó mutatókat határoztuk meg tantárgyanként a vizsgált időszakok teszteredményeire. A negyedéves átlagok összehasonlításával képet kaphatunk arról, hogy a negyedéves tesztek kitöltő hallgatói eredmények hogyan változtak az évek során. Ezek után az egyes évek átlagának alakulását néztük negyedévtől eltekintve tárgyanként, majd a számítások végén összevetettük kurzusonként a kapott eredményeket a hallgatói létszám alakulásával. Az elvégzett számítások során kapott grafikonok, táblázatok mellett érdekelt minket az is, hogy a tesztkörnyezetből exportált adatok hogyan viselkednek egy adatvizualizációt támogató szoftver környezetében. Így néhány adatsort kiemeltünk és a megfelelő adatstruktúra kiépítését követően, beolvastuk őket a Tableau nevű szoftverbe. A leíró statisztika eszköztárának alkalmazásával kapott értékek összesítésével képet kaphatunk az elemzett négy tárgy eredményeinek egymás közötti megoszlásáról, a tárgyak teljesíthetőségéről is.

Méréseink során vizsgálat alá vetettük azt is, hogy létezik-e kapcsolat az évenként elért átlageredmények között, azaz összehasonlítottuk több sokaság várható értékét. A vizsgálatot az Információmenedzsment tárgy eredményeivel végeztük el, mert itt hét évnyi adat állt a rendelkezésünkre. A hipotézisvizsgálat egyik módszerét, a varianciaanalízist használtuk az eredmény megállapításához. Külön figyelmet szenteltünk annak, hogy megvizsgáltuk tárgyanként, illetve minden rendelkezésre álló tesztadatot egyben kezelve a találati arányra számított módusz, illetve medián értékét. Ezekhez a számításokhoz nem a szokásos beépített módusz kiszámítására irányuló Excel függvényeket alkalmaztuk, ugyanis egy-egy ilyen érték segítségével nem tudtuk volna az adatok közötti összefüggéseket leírni, nem jutottunk volna többletinformációhoz az adatsorokról. Így például az egyes tesztsorokhoz tartozó módusz ese-

tében első körben meghatároztuk azt, hogy hány darab osztályköz szükséges az adatok besorolásához és azokhoz milyen alsó/felső határ tartozik, majd a részletes teszteredményeket egy-egy kategóriához rendeltük hozzá, ezzel megkapva, hogy mely osztályhoz tartozik a legtöbb mérési elem.

Kutatómunkánk egyik fő célja az elemzett tárgyak vizsgakérdéseinek minőségi osztályozása, hogy a kapott értékek alapján meg tudjuk határozni, hogy képesek-e a vizsgált kérdések a hallgatói tudást hatékonyan mérni. A kérdések minőségét egyrészt azok nehézsége határozza meg. Egy kérdés nehézségét az alapján mérhetjük, hogy a hallgatók milyen arányban válaszolták meg helyesen. Ehhez szükségünk volt a kérdések felmerülésének és a helyes találatának számára, amelyet az általunk fejlesztett program segítségével nyertünk ki. Az így alkotott találati arány – helyes találatok száma osztva a felmerülések számával – segítségével meghatároztuk, hogy mely kérdések tartoznak a nehezebb, illetve a könnyebb kategóriákba. A kapott eredményeket, még ha valós, megtörtént teszteredményekből kaptuk is, nem tekinthetjük 100%-ig biztosnak, hiszen nem garantált, hogy a következő években történő számonkérések során a hallgatók ugyanilyen arányban fognak jól vagy rosszul válaszolni a kérdésekre, és ugyanazt a találati arányt érik el. Így meg kellett állapítani azokat az intervallumokat, amelyekben mozoghat a találati arány értéke nehéz, illetve könnyű kérdések esetén. Az intervallumokat normális eloszlással, 5%-os szignifikancia szinten határoztuk meg.

Az alsó határt, mely a nehéz kérdéseket jelöli, az alapján adtuk meg, hogy az adott kérdést milyen valószínűséggel találnák el tanulás nélkül, tippeléssel a hallgatók, vagyis mennyinek kell legalább lennie a találati aránynak. Ezt az értéket a válaszok száma határozza meg, hiszen négy válaszlehetőség közül az egyetlen helyes választ 25% esélyünk van eltalálni, 2 válaszlehetőség esetén pedig 50% esélyünk van, hogy a jót választjuk.

A teszteredményekben nem szerepelt, hogy mennyi lehetőség közül lehet választani egy adott kérdésnél, így ezt nekünk kellett a kérdéskatalógusok alapján meghatározni. A kérdések nagy száma miatt kézzel nem volt érdemes hozzáfogni a számoláshoz, így Visual Basic for Applications (VBA kód) segítségével írtunk egy makrót, amely végigfutott a kérdéskatalógus összes celláján és a kérdés azonosítója mellé beírta a válaszlehetőségek számát, továbbá a válaszlehetőségeket a kérdés szövege mellé helyezte egy sorba, majd az üres sorokat feltöltötte a kérdések adataival. Ezt elvégeztük a négy elemzett tárgy esetében számonkérési periódusonként, ahol releváns volt ott az első negyedéves és második negyedéves kérdéseket is rendeztük a későbbi felhasználás, értékmeghatározás céljából pedig összemásoltuk egy munkalapra őket.

A felső határt 80%-nak választottuk, vagyis ha egy kérdést 80%-nál többen válaszoltak meg helyesen, akkor azt a kérdést túl könnyűnek ítéljük meg. Az ideális nehézségű kérdés így 25% (a „tippelési” lehetőség) és 80% közötti találati aránnyal rendelkezik.

| KérdésID | Felmerült | Eltalálta | Arány | Alsó határ | Felső határ | Tipp | Túl egyszerű | Túl nehéz |
|-------------|-----------|-----------|-------|------------|-------------|------|--------------|-----------|
| 1JAS-735LFG | 99 | 6 | 0.061 | 0.014 | 0.108 | 0.25 | | - |
| ZSZO-6ZSEBD | 143 | 133 | 0.930 | 0.888 | 0.972 | 0.25 | + | |
| 1VAS-734DYY | 168 | 166 | 0.988 | 0.972 | 1.004 | 0.25 | + | |
| 1VAS-734DZN | 165 | 157 | 0.952 | 0.919 | 0.984 | 0.25 | + | |
| 1VAS-734E2B | 184 | 31 | 0.168 | 0.114 | 0.223 | 0.25 | | - |
| 3SZO-7CUT5Y | 177 | 146 | 0.825 | 0.769 | 0.881 | 0.25 | | |
| 3SZO-7CUT6F | 191 | 172 | 0.901 | 0.858 | 0.943 | 0.25 | + | |
| 3SZO-7CUT6V | 171 | 119 | 0.696 | 0.627 | 0.765 | 0.25 | | |

2. ábra: Kérdés nehézség meghatározása (részlet)

Forrás: Saját szerkesztés

Az 5%-os szignifikancia szinttel kapott intervallum értékeket és határokat vetettük össze egymással, és ez által határoztuk meg az egyes kérdések nehézségét.

Gyakorlatban ez a következőképpen néz ki: egy 6%-os találati arányú kérdésnek a konfidencia intervallumát határoztuk meg 5% szignifikancia szinten, melynek alsó értéke 1,4%, felső értéke 10,8% lett. Az itt bemutatott kérdés négy válasz lehetőséggel rendelkezett és mivel csak 1 jó válasz van, így az alsó határ 25% lett. A kérdés teljes konfidencia intervalluma a határ alatt áll, így a kérdést rendkívül nehéznek ítéltük meg. A bemutatott adatok a következő kérdésre vonatkoznak:

Mely területen támogatja a menedzsmentet a konfigurációmenedzsment (1JAS-735LFG):

- *terméktulajdonos (aki felelős a projekttermékért),*
- eseménynapló rendszeres felülvizsgálatát,
- a projekttel kapcsolatos lekérdezéseket,
- a tevékenységek dokumentálását és ellenőrzését.

A következő kérdést, az előző bemutatott példával ellentétben a 143 felmerülés alatt 133-szor válaszolták meg helyesen (ZSZO-6ZSEBD):

Az információrendszer...

- információt gyűjt, szolgáltat, létrehoz, tárol, szétválogat, használ és eloszt;
- minden esetben számítógépekre épül;
- nem célorientált, rugalmasan használható a szervezeti célokra;
- elsősorban a vezetői/irányítási munkát, és nem a mindennapi működést támogatja.

| KérdésID | Felmerült | Eltalálta | Arány | Alsó határ | Felső határ | Tipp | Túl egyszerű | Túl nehéz |
|-------------|-----------|-----------|-------|------------|-------------|------|--------------|-----------|
| 1JAS-735LFG | 99 | 6 | 0.061 | 0.014 | 0.108 | 0.25 | | - |
| ZSZO-6ZSEBD | 143 | 133 | 0.930 | 0.888 | 0.972 | 0.25 | + | |

3. ábra: Példa nehéz és könnyű kérdésre

Forrás: Saját szerkesztés

Egy kérdésről nem csak az állapítható meg, hogy könnyű vagy nehéz. Az item analízis egyik mutatója, az item elkülönítésmutató további információkkal szolgálhat a kérdésről. A mutató megvizsgálja, hogy az adott kérdést helyesen megválaszolók közül milyen arányban vannak jó, közepes és rossz hallgatók. Ha csak jó teljesítményt nyújtó hallgató találta el a kérdést, akkor a kérdés rossz. Ennek ellenőrzése szintén hasznos segítséget nyújthat a kérdések fejlesztésében.

A kérdések minősítésével megállapítottuk, hogy elsősorban mely kérdésekkel szükséges foglalkozni. Annak érdekében, hogy még részletesebb információt kapjunk a kérdésekről, a programba újabb lekérdezést implementáltunk, melynek segítségével megállapítható, hogy az adott kérdés válaszlehetőségei közül melyikre hány jelölés érkezett.

Feleletválasztós tesztek kitöltése során, hasznos módszer lehet, ha nem tudjuk elsőre a választ, hogy első körben kizárunk a lehetőségek közül párat. Ha biztosan ki tudunk zárni egy lehetőséget, akkor már eggyel kevesebb lehetőség közül kell választanunk, amellyel a kockázatunk csökken és az esélyünk nő a helyes válasz eltalálására. A tudás, felkészültség mérésének szempontjából fontos, hogy a hallgató ne kitotózza a helyes választ, hanem határozottan tudja azt, hiszen akkor mélyültek el igazán a félév alatt tanult ismeretek.

Ha egy adott kérdés adott válaszai közül hármát senki sem jelölt be, akkor érdemes megvizsgálni, hogy miért nem. Lehetséges, hogy annyira sokszor hallották adott kontextusban a kifejezést, amelyre rákérdeznek, hogy automatikusan azt a választ jelölik be, amelyik helyes. Ebben az esetben négy válaszlehetőség esetén háromra 0 közeli jelölés érkezik. Viszont az is lehetséges, hogy az adott válaszlehetőséget a hallgatók túl egyszerűen zár-

ták ki, mert nincs semmi kapcsolata a tananyaggal, vagy hibás logikailag a válasz, mint a következő esetben:

Az informatika gazdasági alkalmazásának történetét a „három kor” fejlődéstörténeti modell írja le a következő fázisok szerint (1SZO-6VUCXH):

- *Adatfeldolgozó rendszerek kora, vezetői információrendszerek kora, stratégiai rendszerek kora.*
- A kísérletezés, az elterjedés, a kontroll és végül az érettség kora (az IT stratégiai alkalmazása).
- Stand-alone rendszerek kora (nagygépek), terminálos rendszerek kora.
- WAN-, MAN-, LAN-korszak és a jelenlegi Internetes korszak ismeretes.

Könnyen látható a megoldás bármilyen előzetes tanulás nélkül is, ezt bizonyítja az is, hogy a 140 felmerüléséből 139-en eltalálták. A kérdésben is szerepel a “három kor”, amely arra utal, hogy három darab kifejezés lesz a válaszokban. Ha megnézzük jobban, akkor láthatjuk, hogy csak az első válasz esetében áll fent, hogy „három kor” van felsorolva, a többi esetben csak kettő vagy négy.

| KérdésID | Helyes válasz | Megoszlás | | | | | | | |
|-------------|---------------|-----------|------|------|------|-------|------|------|------|
| | | 1-es | 2-es | 3-as | 4-es | 1-es | 2-es | 3-as | 4-es |
| 1SZO-6VUCXH | 1 | 176 | 2 | 2 | 1 | 97.2% | 1.1% | 1.1% | 0.6% |

4. ábra: Hibás válaszlehetőségek

Forrás: Saját szerkesztés

Az áttekinthetőség és egyértelműség céljából százalék, megoszlás formájában dolgoztunk a válaszlehetőségekre érkezett jelölésekkel. Fontos, hogy nem csak az a válaszlehetőség érdemel figyelmet, melyet 0 ember jelölt, hanem egy bizonyos százalék alatti érték is gyanús lehet.

A következő példában viszonylag sok jelölés érkezett három válaszlehetőségre is a felmerült 138 esetben, viszont az egyik válaszlehetőséget csak egy hallgató jelölte be (1BEE-735J4V):

IT mint szervezetátalakító tényező vizsgálatakor:

- a hangsúly az információtechnológia műszaki részletein van.
- *a hangsúly azon van, hogy az információtechnológia képes-e támogatni az adott üzleti folyamatot.*
- a hangsúly azon van, hogy az információtechnológia képes-e támogatni a szervezet egészét.
- a hangsúly azon van, hogy az információtechnológia képes-e radikálisan átalakítani az üzleti folyamatokat.

| KérdésID | Helyes válasz | Megoszlás | | | | | | | |
|-------------|---------------|-----------|------|------|------|------|-------|-------|-------|
| | | 1-es | 2-es | 3-as | 4-es | 1-es | 2-es | 3-as | 4-es |
| 1BEE-735J4V | 2 | 1 | 19 | 66 | 52 | 0.7% | 13.8% | 47.8% | 37.7% |

5. ábra: Hibás válaszlehetőség

Forrás: Saját szerkesztés

Tanulmányaink során azt tapasztaltuk, hogy olykor célszerű a leghosszabb és legbonyolultabbnak tűnő választ bejelölni, ha nem tudunk dönteni biztosan a lehetőségek közül. Körüljártuk ezt a feltételezést is, hogy valójában számíthat-e a diák arra, hogy pontot szerez, ha a leghosszabb választ jelöli, továbbá azt is vizsgáltuk, hogy a diákok alkalmazzák-e ezt a stratégiát a tesztek kitöltése során.

A vizsgált tantárgyak tesztrendszereihez tartozó kérdéskatalógusok kérdéseit oktatói segítséggel minden vizsgált tárgy esetében a számonkérés alapját képező tananyagok egy-egy témaköréhez, fejezetéhez rendeltük. A kategorizálás segítségével választ kaptunk arra, hogy a tesztrendszer témaköréi, kérdései hogyan épülnek fel, mely témaköröket érdemes további kérdésekkel bővíteni, melyek azok a részek, amelyek az elvárt kérdésszámtól jóval többet vagy éppen kevesebbet tartalmaznak. Elemzéseink során hangsúlyt fektettünk arra, hogy feltárjuk azokat az okokat, amelyek a vizsgált időszakok alatt a kérdéskatalógusokban bekövetkező változásokat, változtatásokat idézték elő.

A kategorizálás után különböző számításokat végeztünk el a kibővített adatokon. Megvizsgáltuk, hogy mely témakörök tekinthetők a legnehezebb illetve a legkönnyebb részeknek az adott tantárgyi vizsgasorokban, azaz összegeztük fejezetenként, témakörönként a kérdések felmerülésének és helyes találatának számát, majd arányosítottuk őket egymáshoz.

Végül, kíváncsiak voltunk arra is, hogy a tantárgyakon belül témakörökre lebontva hogyan viszonyul egymáshoz az adott témakörhöz kapcsolódó kérdések száma és azok tényleges előfordulása a vizsgált időszak alatt. Ezzel azt kívántuk mérni, hogy mutatnak-e egyenletességet a rendelkezésre álló adatok az egyes számonkérési időszakok során vagy sem.

Következtetések (összefoglalás):

A dolgozatban a Budapesti Corvinus Egyetem Információrendszerek Tanszékének egyik meghatározó tesztkörnyezetében futó, feleletválasztós tesztek, teszt sorok elemzéséhez és kiértékeléséhez készítettünk el egy keretrendszert.

A 2009 és 2016 közötti Információmenedzsment, a 2014 és 2016 közötti Üzleti informatika és Üzleti intelligencia, illetve a 2015 és 2016 közötti Integrált rendszerek tárgyak számonkérési rendszereinek adatai és a vizsgáztatás során felmerülő hallgatói tesztteredmények képezték számításaink alapjait. A statisztikai és matematikai elemzéseket kiegészítettük egy program fejlesztésével, mely képes a tesztkörnyezetből exportált adatokat kezelhető, elemzésre alkalmas formába hozni. Az idő előrehaladásával olyan funkciókat implementáltunk az alkalmazásba, amelyek képesek hozzájárulni ahhoz, hogy gyors visszajelzés álljon az oktatók rendelkezésére a teszt sor hatékonyságáról, eredményességéről a számonkérést követően.

Az alábbi ábra foglalja össze az általunk elvégzett méréseket, elemzéseket:



6. ábra: Összefoglaló

Forrás: Saját szerkesztés

A leíró statisztika eszközeit felhasználva a kapott eredményeink alapján a főbb megállításaink közé tartozik az, hogy a hallgatói létszám növekedésével a rosszabb teszt eredmények részaránya is növekszik egy adott kurzusnál, viszont a jó eredményeké szinte változatlan marad.

A kiválasztott item mutatók értékeinek megadásával és az egyes kérdésekhez tartozó válaszok és azok adatainak elemzésével arra jutottunk, hogy a kérdések és válaszok létrehozásánál érdemes figyelni és nagyobb hangsúlyt fektetni a válaszlehetőségek kidolgozására, ha a tényleges hallgatói tudás felmérésére törekszünk, hogy ne legyenek hibásan felépített kérdések, illetve logikailag kizárható válaszlehetőségek.

Az elvégzett számítások alapján érdemes lehet tantárgyanként átnézni és közelebbről megvizsgálni a nehéz vagy könnyű kérdésnek kategorizált itemeket és azok válaszait, ugyanis összesen 201 olyan kérdés szerepel mostanra a négy tárgy tesztrendszerében a megközelítőleg 1200 kérdésből álló összevont kérdéskatalógusban, melyet 5% szignifikancia szint mellett a hallgatók legalább 80%-a eltalált, illetve 47 olyan kérdés, melyet a hallgatók 25%-a sem válaszolt meg helyesen, így ezek a kérdések nem képesek megfelelően differenciálni a hallgatói tudást.

Az adott kurzus tesztrendszeréhez kapcsolódó kérdéseket csoportosítottuk az előírt elméleti tananyag egy-egy témaköre, fejezete alapján. Ezt követően számításokat végeztünk el, hogy megtudjuk, hogyan épülnek fel az egyes tantárgyi kérdéskatalógusok, melyek a legnehezebben megválaszolható tananyagrészek, hogy a kapott eredményekből következtetni tudjunk arra, mely témakörökre érdemesebb több időt szánni az oktatás során. Emellett a statisztikai hipotézisvizsgálat segítségével megvizsgáltuk, hogy az egyes kérdéskatalógusokban szereplő kategorizált kérdésszámok mutatnak-e egyenletességet a tényleges előfordulásukhoz képest.

Arra a következtetésre jutottunk, hogy a feleletválasztós tesztrendszerek bevezetése óta számos változás ment végbe, ami felborította a kérdéskatalógusok felépítését, így az egyenletesség az adatok között csak bizonyos kiugró értékek elhagyásával mutatható ki, vagy nem mutatható ki. Ezt erősítette az is, hogy a kérdések rendszerezése során bebizonyosodott, hogy vannak olyan kérdések, amelyek az évek alatt teljesen eltűntek a rendszerből.

Előfordul az is, hogy a kérdéskatalógusban a fejezetenkénti kérdésszám sem azonos, előfordulnak olyan fejezetek, amelyek alig tartalmaznak kérdést, és olyanok is, amelyek kérdésszáma jóval magasabb az elvárt átlagértéknél. Megállapítható, hogy a kérdéskatalógusokban bekövetkező változások, illetve az eltérő kérdésszámú kategóriák azt is eredményezhetik, hogy bizonyos fejezeteket a hallgatók jóval nehezebbnek ítélnék meg (általában ezek az alacsonyabb kérdésszámú témakörök), szemben a magasabb kérdésszámmal rendelkező tananyagrészekkel, amelyek számonkérésénél a hallgatók jobban teljesítenek.

Köszönetnyilvánítás

Köszönettel tartozunk dr. Vas Réka Franciska (Budapesti Corvinus Egyetem Információrendszerek tanszék) tanárnőnek, hogy munkánkat segítette, a megfelelő úton tartott minket, továbbá köszönettel tartozunk Balázné Mócsai Andrea (Budapesti Corvinus Egyetem Statisztika Tanszék) tanárnőnek, hogy munkánkat statisztikai szemmel vizsgálva folyamatosan segítette és végigkísérte. Az érdem közös, mert munkájukkal nagy mértékben járultak hozzá a tanulmányunk megvalósulásához.

Köszönettel tartozunk a Budapesti Corvinus Egyetemnek és az Emberi Erőforrások Minisztériumának, hogy az ÚNKP-16-1 kódszámú Új Nemzeti Kiválóság Programjukon keresztül 2016 szeptemberétől 2017 júniusáig támogatták szakmai tevékenységünket és kutatási munkánkat.

Irodalomjegyzék

- Csapó Benő (2002): Az iskolai tudás. Osiris Kiadó, Budapest
- Kontra József (2011): A pedagógiai kutatások módszertana. Kaposvári Egyetem, Kaposvár (letöltve: 2016.03.18.)
http://janus.ttk.pte.hu/tamop/kaposvari_anyag/kontra_jozsef/ch04s03.html
- Fercsik János (2014): Pedagometria fogalma. Kislexikon.hu (letöltve: 2017.03.10.)
<http://www.kislexikon.hu/pedagometria.html>
- Horváth György (1991): Az értelem mérése. Tankönyvkiadó, Budapest (letöltve: 2016.11.25.)
- Horváth György (1993): Bevezetés a tesztelméletbe. Keraban Kiadó, Budapest
http://miau.gau.hu/miau/159/tdk/ai_Varga_Viktor.pdf
- Mehrens W. A. - Lehmann, I. J. (1973). Measurement and Evaluation in Education and Psychology. New York: Holt, Rinehart and Winston, pp. 333-334., [DOI: 10.1177/001316447303300444](https://doi.org/10.1177/001316447303300444)
- Murphy K. R. - Davidshofer (1994): Psychological testing. C. O. N. J., Prentice Hall International Englewood Cliffs, [DOI: 10.1002/9781118785317.weom050161](https://doi.org/10.1002/9781118785317.weom050161)
- Nagy József (1975): A témazáró tesztek reliabilitása és validitása. Acta Universitatis Szegediensis de A. J. Nominatae, Sectio Paedagogica et Psychologica, Series Specifica, Szeged
- SAKAPKMM (2010): Segédlet A Kompetencia Alapú Pedagógus Képzés Módszertani Megújulásához. TÁMOP-4.1.2/B projekt, Győr-Moson-Sopron Megyei Pedagógia Intézet (letöltve: 2017.03.11.)
http://pszk.nyime.hu/tamop412b/meres_ertekeles/index.html